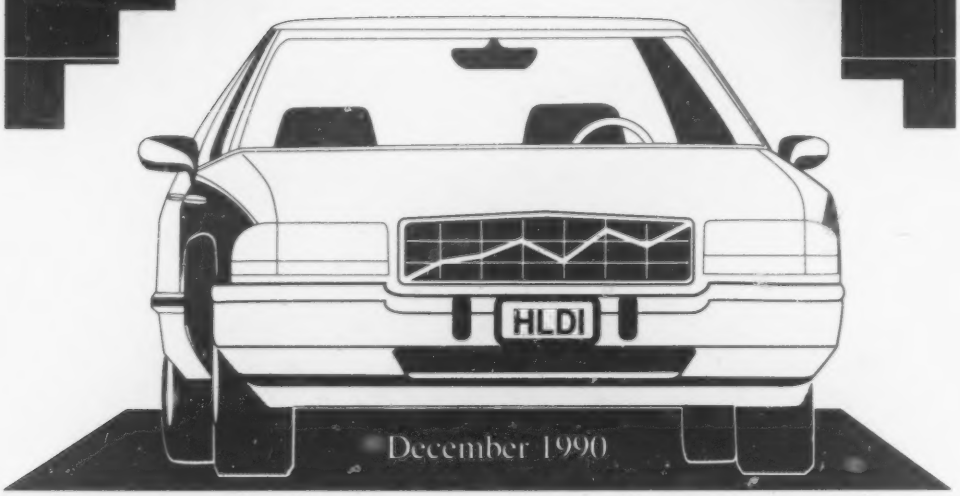
A5200-6F 1/91

TECHNICAL APPENDIX



December 1990

HIGHWAY LOSS DATA INSTITUTE

Permission to quote from or to reproduce material from this publication is granted when due acknowledgement is made.
Highway Loss Data Institute -- 1005 North Glebe Road -- Arlington, Virginia 22201
January 1991 -- Copyright 1991 HLDI

GLOSSARY

Average Loss Payment per Claim -- The total of all loss payments made for the claims for a group of vehicles, divided by the number of claims paid.

Average Loss Payment per Insured Vehicle Year -- For a group of vehicles, the product of Claim Frequency and Average Loss Payment per Claim, expressed as dollars per insured vehicle year. It should be noted that this definition differs from the commonly used insurance term, pure premium, but yields similar results.

Claim Frequency -- The number of claims for a group of vehicles, divided by the exposure for that group; expressed as claims per one hundred or one thousand insured vehicle years.

Collision Coverage -- As defined in insurance contracts, the coverage under which people insure their own vehicle against loss caused by collision.

Deductible Amount -- That portion of the collision or theft damage cost borne by the policyholder. For computing results, deductible amounts are split into two groups, a low deductible group and a high deductible group.

Exposure -- The time interval an individual vehicle is insured. Exposure for a group of vehicles is expressed in units of insured vehicle years.

First-Party Coverage -- An insurance coverage under which the policyholder collects compensation for his losses from his own insurer regardless of fault.

Loss Payment -- That portion of the insurance loss cost borne by the insurer; in general, the total loss cost minus the deductible amount.

No-Fault Auto Insurance -- An auto insurance plan in which each driver/owner accepts financial responsibility for some or all losses sustained by himself, pedestrians hit by him, and occupants of his own vehicle in return for which he enjoys some immunity from liability for losses to third-party persons.

Operator Age Group -- The factor that distinguishes those vehicles with a youthful rated driver from those vehicles without a youthful rated driver (see Youthful Operator).

Personal Injury Protection (PIP) Coverage -- First-party no-fault coverage in which an insurer pays, within the specified limits, the medical, hospital, and other expenses of the insured, others in his vehicle, and pedestrians struck by him.

Rated Driver -- The driver who, for insurance purposes, is considered to represent the greatest loss potential for the insured vehicles.

Theft Coverage -- The coverage under which people insure their own car against loss caused by partial or total theft.

Youthful Operator -- All males under 25 years of age and all unmarried females under 25 years of age.

DATA SOURCES

Data Supplying Companies

The Highway Loss Data Institute (HLDI) receives claims and coverage information from the following automobile insurance companies:

Aetna
Allstate
Fireman's Fund
GEICO
Kemper
Liberty Mutual
Nationwide
Prudential
Safeco
State Farm
Travelers
USAA

Data are collected for collision and theft coverages for the three most recent model years and for personal injury protection coverage for the five most recent model years. Assigned risk and fleet policies are not included in the HLDI data base. All data for the state of Massachusetts are excluded as well. Only payments for damage to insured vehicles are included, that is monies recovered by the companies from either the salvage of wrecked vehicles or through the subrogation process are excluded.

Insurance Policy Data

Insurance policy data, as distinct from claim data, describe the characteristics of an insured vehicle, including its make and series and the scope of its insurance coverage. Policy data reported to HLDI include the following basic information:

- Type of Coverage
- Vehicle Identification Number (VIN)
- Deductible Amount
- Rated Driver Characteristics
- Date and Nature of Relevant Changes in Coverage
- Geographic Location

Insurance Claim Data

Insurance claim data report the characteristics of an insured vehicle for which a claim is made, the initiation of the claim, and/or the corresponding payment. Claim data reported to HLDI include the following basic information:

- Type of Coverage
- Vehicle Identification Number (VIN)
- Crash Date
- Loss Payment Amount

The dollar amounts reported represent loss payments made to, or on behalf of, the policyholder for damage to the insured vehicle.

Vehicle Description and Identification

The vehicle identification numbers (VINs) of most new automobiles sold in the United States are individually unique numbers that contain, in coded form, detailed information as to each vehicle's make, series, and other distinguishing characteristics.

The specific vehicle types for which results are presented in this report were derived from the VINs of the individual automobiles.

Only data for vehicles with valid VINs were used for this report. Each VIN was decoded into a set of specific vehicle characteristics. For the purposes of this report the vehicle characteristics of primary concern were make, series, vehicle size class, and vehicle body style. The use of VINs also permits analyses of other very specific vehicle characteristics, such as engine type and horsepower.

DATA TYPES

Exposure -- All Coverages

Exposure, in insured vehicle years, is computed for each individual vehicle from the coverage data. The total number of insured vehicle years for each vehicle series was then obtained by accumulating the exposure for all of the individual vehicles of that vehicle series.

Claim Frequency -- All Coverages

Both pending and paid claims for an individual vehicle are first matched with its corresponding coverage data to ensure that the claim had occurred within an appropriate period of insurance coverage for that vehicle. When multiple claims are made for the same crash, they are treated as a single claim in the calculations of claim frequency. Claims that are settled without payment, as well as those paid even though the damage was less than the deductible amount, are excluded from the computations. All remaining claims for each vehicle series are used, together with the corresponding exposure, to compute the claim frequency. For collision coverage, claim frequency is expressed as the number of claims per 100 insured vehicle years. For injury and theft coverages, claim frequencies are expressed as the number of claims per 1000 insured vehicle years due to the much less frequent nature of these claims.

Average Loss Payment Per Claim -- Collision and Theft Coverages

Paid claims for each vehicle series are summed and divided by the number of claims to produce the average loss payment per claim for that series. Claims that are settled without payment, as well as those paid even though the damage was less than the deductible amount, are excluded from the computations. Where multiple claims and/or multiple payments are made for the same crash, they are treated as a single claim and/or single payment in the calculation of the average loss payment per claim.

Average Loss Payment Per Insured Vehicle Year -- Collision and Theft Coverages

The average loss payment per insured vehicle year for each vehicle series is obtained by multiplying the claim frequency per 100 insured vehicle years (1000 insured vehicle years for theft results) by the average loss payment per claim, and dividing the result by 100 for collision results (1000 for theft results).

METHODS

Vehicle Classifications

Passenger Cars

There are five major body style groupings of passenger cars: regular two-door models, regular four-door models, station wagons/passenger vans, sports models, and luxury models. Sports models include small and midsize two-seater cars, as well as cars with significant high performance features. The luxury models include relatively expensive large and midsize cars that are not classified as sports models. Body styles are divided into size classes based on wheelbase (rounded to the nearest inch) defined as follows:

Small - Vehicles with wheelbases less than or equal to 99 inches.

Midsize - Vehicles with wheelbases greater than 99 inches and less than or equal to 109 inches.

Large - Vehicles with wheelbases greater than 109 inches.

Vans

Vans are fully enclosed vehicles with either no hood or a very short hood, in which the engine is placed at least 50 percent behind the windshield. The driver's position is well forward, within the front 25 percent of the wheelbase. These vehicles are designed for cargo transport and are equipped with a right side cargo access door and rear door(s). The cargo area is not separated from the passenger area, both are enclosed under the same roof. Some versions of these vans are equipped with additional seats in the cargo area and usually have additional side windows as well. There are no size classes for vans.

Pickups

Pickups are cargo-carrying vehicles, usually on a truck chassis, with an enclosed cab and a separate open cargo area. The open cargo area generally is a box with raised sides with an opening tailgate. Pickups are divided into two size classes based on curb weight.

Small - Pickups weighing less than 3,500 pounds.

Standard - Pickups weighing 3,500 pounds or more.

Utility Vehicles

Utility vehicles are built on heavy-duty chassis capable of off-road use. They are of conventional construction with the driver behind the engine. The cargo and passenger areas are integral with the driver area. Many of these vehicles are equipped with soft or removable tops. Frequently these vehicles are equipped with four-wheel drive. Utility vehicles are divided into three size classes based on wheelbase.

Small - Vehicles with wheelbases less than 100 inches.

Intermediate - Vehicles with wheelbases from 100 to 120 inches.

Large - Vehicles with wheelbases greater than 120 inches.

Threshold for Reporting Results

Insured vehicle years of exposure forms the measure of sample size for HLDI results. In general the targeted minimum reliability standard for presentation of HLDI frequency results is that the estimated value falls within $\pm$ 20 percent of the true value 90 percent of the time. More reliable results are obtained for models with larger amounts of exposure. For collision coverage results to be

presented, an individual vehicle series must have at least 1000 insured vehicle years of exposure. For theft coverage the reporting threshold is 8,000 insured vehicle years of exposure. Injury results are presented for individual vehicle series with more than 5,000 insured vehicle years of exposure. In addition for injury results, the different thresholds which are used to determine whether the claim frequency results for medical losses exceeding specified dollar are published are given in the following table.

Injury Losses	*Reporting Threshold*
>$250	10,000 insured vehicle years
>$500	15,000 insured vehicle years
>$1,000	25,000 insured vehicle years

The relative injury loss experience of particular vehicles does not change substantially from model year to model year providing that the basic design of the vehicle and the type of restraint system remains essentially unchanged. For this reason and because substantial amounts of exposure are needed to provide reliable injury claim frequency results, data for the most current three model years are combined for those vehicles with essentially unchanged designs.

The aggregated model year result (R) for each vehicle series (i) is obtained by weighting the relative claim frequency for each model year (j) by the applicable exposure for each model year; that is,

$$R_i = [\Sigma (exp_{ij} (f_{ij} / F_j)) / (\Sigma exp_{ij})] \times 100,$$

where

exp_{ij} = exposure for vehicle i for model year j,

f_{ij} = standardized claim frequency for vehicle i for model year j, and

F_j = standardized claim frequency for all passenger cars for model year j.

EXAMPLE:

	1990	1989	1988
Exposure (exp)	11,994	22,301	30,458
Claim Frequency (f)	23.6	19.9	17.6
Claim Frequency (F) -- all	20.2	20.8	20.0

Result = [11,994 (23.6 / 20.2) + 22,301 (19.9 / 20.8)+ 30,458 (17.6 / 20.0)] / [11,994 + 22,301 + 30,458] x 100 = 96.

Standardization

It is well known that loss experience can vary substantially in relation to certain nonvehicle factors. Two important such factors are the deductible amount of the coverage and the operator age group; both claim frequency and average loss payments per claim vary with these factors. In the case of the operator age group, youthful operators generally generate higher insurance losses than older drivers. Thus, if compared vehicles have different proportions of youthful operators, this difference, if not taken into consideration, would be sufficient to bias the comparison.

For collision coverage, results obtained from the two deductible categories (less than or equal to $200 and greater than $200) also vary. For example, average loss payments per claim for the higher deductibles are greater than those for the lower deductibles. Again, if compared vehicles have different proportions of lower deductible coverages, this difference, if not taken into consideration, would also be sufficient to bias the comparison. A similar approach is necessary for theft coverages with the lower deductible category including deductibles less than or equal to $50 and the higher category being those greater than $50. No deductible is available for Personal Injury Protection coverages.

To minimize any biases in the comparison of results that could arise because of differences attributable to variations among different vehicles in deductible amount and operator age groups, collision and theft results are first adjusted or standardized so as to equalize the effects of the two nonvehicle factors in the standardized results. In this procedure, a weighted average of the actual results for each combination of deductible amount and operator age group is computed using the standardization weights shown in the table below. Because the same weights are always used, the effects due to the nonvehicle factors are present in equal amounts in the standardized results. Therefore, to the extent that the age distributions within each operator age group and the deductible distributions within each deductible group do not vary substantially for different vehicles, these effects can no longer bias comparisons of results. Standardization of injury results takes place in exactly the same manner using only the combined operator age weights as there is no deductible in Personal Injury Protection coverage.

The standardization procedure employed is widely used in health statistics.* The basic principle of the method is the introduction of a selected population with a standard distribution across the combinations of deductible amount and operator age group as represented by the standardization weights in the following tables.

Standardization Weights: Collision Coverage

Deductible Category	Operator Age Group: Youthful Operator	No Youthful Operator	Combined
≤$200	0.05	0.50	0.55
>$200	0.05	0.40	0.45
Combined	0.10	0.90	1.00

Standardization Weights: Theft Coverage

Deductible Amount	Operator Age Group: Youthful Operator	No Youthful Operator	Combined
≤$50	0.05	0.50	0.55
>$50	0.05	0.40	0.45
Combined	0.10	0.90	1.00

Standardization Weights: Personal Injury Protection Coverage

Deductible Amount	Operator Age Group: Youthful Operator	No Youthful Operator	Combined
None	0.10	0.90	1.00

Mathematical Procedures

The standardization procedure can be represented mathematically as follows:

Deductible Amount	Operator Age Group: Youthful Operator	No Youthful Operator	Combined
Lower	X_{11}	X_{12}	$X_{1.}$
Higher	X_{21}	X_{22}	$X_{2.}$
Combined	$X_{.1}$	$X_{.2}$	$X_{..}$

* For example, See Armitage, P. *Statistical Methods in Medical Research*, Wiley, New York (1971).

where,

- X_{ij} represents a computed claim frequency for a particular vehicle series with respect to the deductible amount i and age group j, and
- i = 1 for lower category deductibles and 2 for higher category deductibles,
- j = 1 for youthful operator group and 2 for no youthful operator group,
- $X_{.j}$ represents the computed claim frequency for all deductibles combined for age group j,
- $X_{i.}$ represents the computed claim frequency for both operator age groups combined for deductible amount group i, and,
- $X_{..}$ represents the computed claim frequency for all deductibles and both age groups combined.

Now, if weights are defined as follows:

Deductible Amount	Youthful Operator	No Youthful Operator	Combined
Lower	W_{11}	W_{12}	$W_{11} + W_{12}$
Higher	W_{21}	W_{22}	$W_{21} + W_{22}$
Combined	$W_{11} + W_{21}$	$W_{12} + W_{22}$	1

where $W_{11} + W_{21} + W_{12} + W_{22} = 1$,

then the standardized claim frequency for the particular vehicle series is:

$$W_{11}X_{11} + W_{21}X_{21} + W_{12}X_{12} + W_{22}X_{22},$$

and, if A_{ij} represents the computed average loss payment per claim for a particular vehicle series, then the standardized average loss payment per claim is:

$$\frac{W_{11}X_{11}A_{11} + W_{21}X_{21}A_{21} + W_{12}X_{12}A_{12} + W_{22}X_{22}A_{22}}{W_{11}X_{11} + W_{21}X_{21} + W_{12}X_{12} + W_{22}X_{22}}$$

The standardized average loss payment per insured vehicle year is obtained by multiplying the standardized claim frequency by the standardized average loss payment and dividing by 100 in the case of collision coverage or 1000 for theft coverage.

Similarly, it is possible to standardize within a deductible category or within an operator age group by using the appropriate combined weights. The standardized claim frequency for youthful operators, for example, would be:

$$(W_{11} + W_{12})X_{11} + (W_{21} + W_{22})X_{21},$$

and the standardized claim frequency for higher category deductibles would be:

$$(W_{11} + W_{21})X_{21} + (W_{12} + W_{22})X_{22}.$$

These latter standardizations, using the combined weights, enable valid comparisons to be made between the results for different deductible groups or between the results for different operator age groups without the effects of the other factor biasing the comparisons.

EXAMPLE

For example, assume that for a particular vehicle series, the four claim frequencies are:

$$X_{11} = 16.8 \quad X_{12} = 11.9 \quad X_{21} = 15.8 \quad X_{22} = 10.3$$

and the weights are:

$$W_{11} = 0.05 \quad W_{12} = 0.50 \quad W_{21} = 0.05 \quad W_{22} = 0.40$$

The standardized claim frequency is:

$$(0.05)(16.8) + (0.50)(11.9) + (0.05)(15.8) + (0.40)(10.3) = 11.7.$$

Monthly Exposure Adjustment - Initial Collision Results Only

To compare the results from new model year cars during their initial period of availability with the corresponding results from the previous model year's cars, it is necessary to adjust the latter

results because of differences in the mix of exposure by month that contribute to that result. For the new model year exposure in each successive month increases as more of the new model cars are sold, and therefore there is more exposure from the later months than from the earlier months. For this same period, nearly all of the previous model year cars are on the road so that exposure in each month is essentially the same.

The effects of this difference in exposure can be removed by standardizing the monthly proportions of exposure in the results for the two model years. In the adjustment employed, results for the previous model year cars were adjusted to have the monthly mix of exposure corresponding to the new models. The actual adjustment weights vary by year; the following table gives the monthly weights for the 1990 models as an example.

Monthly Adjustment Weights

Jan.-June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.
0.01	0.01	0.02	0.03	0.06	0.11	0.18	0.26	0.32

Analogous to the earlier discussion of the standardization procedure for age and deductible amount, let

X_m represent a computed claim frequency adjusted for deductible amount and age in month m for a particular vehicle series, and

W_m represent the monthly adjustment weights cited in the preceding table.

Thus, the standardized claim frequency adjusted for monthly exposure is:

$$\Sigma W_m X_m .$$

Similarly, if A_m represents the computed average loss payment per claim adjusted for deductible amount and age in month m for a particular previous model year vehicle series, then the standardized average loss payment per claim is:

$$\frac{\Sigma W_m X_m A_m}{\Sigma W_m X_m} .$$

Thus, using these weights, weighted averages of standardized monthly claim frequency and average loss payment per claim results were computed for previous model year cars. Recall that the standardized average loss payment per insured vehicle year is then obtained by multiplying the standardized claim frequency by the standardized average loss payment per claim and dividing by 100. The results thus obtained are both standardized for age and deductible and adjusted for monthly exposure differences.

An assumption made in computing a standardized and adjusted result is that exposure in each month is independent of the factors deductible amount and operator age group. This assumption is valid because the distributions of exposure by deductible amount and operator age group are approximately the same for each month.

HIGHWAY
LOSS
DATA
INSTITUTE

1005 NORTH GLEBE ROAD, SUITE 800, ARLINGTON, VA 22201